

The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins

A.A. Bukreyev, V.E. Volchkov, V.M. Blinov and S.V. Netesov

NPO 'VECTOR', Institute of Molecular Biology, 633159 Koltsovo, Novosibirsk Region, Russian Federation

Received 26 March 1993

cDNA was synthesized and cloned on the template of the genomic RNA of Marburg virus (strain Popp). Recombinant plasmids with specific cDNA inserts were selected and sequenced. The length of the open reading frame encoding the GP-protein is 681 amino acids. GP-protein is proposed to be an integral membrane protein. Computer-assisted comparison of the deduced amino acid sequence with those of different viruses revealed significant homology with the GP-protein of Ebola virus and with the 'immunosuppressive domain' of the P15E envelope proteins of some oncogenic retroviruses.

Marburg virus; Filoviridae; cDNA; Envelope protein; Immunosuppressive domain

1. INTRODUCTION

Marburg virus causes severe hemorrhagic fever in humans with a mortality rate of 30–35% [1]. The first outbreaks occurred in 1967 in Germany and Yugoslavia when infected monkeys were imported from Africa [1]. Since then some cases of Marburg disease have been reported [2,3]. Together with the other extremely virulent virus, Ebola virus, Marburg virus belongs to the family Filoviridae [4].

The genome of the Marburg virus is a non-segmented RNA strand of negative sense [5]. It encodes seven structural proteins in the following gene order: 3'-NP-VP35-VP40-GP-VP30-VP24-L-5' [6]. All these proteins are produced from monocistronic messenger RNAs which are complementary to the negative-strand genomic RNA [5,6]. The GP-protein is a glycosylated envelope protein and is presented in virus particles as a homotrimer [7]. Identification of the primary structure of the GP-gene and its product is of great importance in understanding the nature of the virus and the mechanisms of its high pathogenicity.

Our first results of cDNA synthesis, cloning and se-

quencing of Marburg virus genome fragment were published in [8]. In this paper we report the synthesis of a cDNA on a genomic RNA template, its cloning, the nucleotide sequence of the GP-gene and its deduced amino acid sequence.

2. MATERIALS AND METHODS

The Popp strain of Marburg virus was isolated in 1967 in Frankfurt, Germany. Virus was obtained from the Byelorussian Institute of Epidemiology and Microbiology (Minsk, Byelorussia). Virus was purified from plasma of infected guinea pigs by gradient ultracentrifugation as described previously [8]. Isolation of genomic RNA was performed as described in [9]. cDNA was synthesized using AMV reverse transcriptase and a random primer. A cDNA-RNA hybrid was cloned into plasmid pBR322 after poly(C) tailing [9]. Specific clones were screened by colony hybridization on nitrocellulose filters with [³²P]ATP-labeled genomic RNA. The primary structure was determined by the Maxam and Gilbert method [10]. A homology search through SWISSPROT or EMBL bank was performed by the QUICK program of the GENBEE package.

3. RESULTS AND DISCUSSION

More than 1,000 hybrid plasmids containing specific viral inserts were screened by cross-hybridization and restriction endonuclease mapping, of which 37 partly overlapping plasmids were sequenced. It was shown that these plasmids contained the complete nucleotide sequence of the Marburg virus genome (more than 19,100 bp). Seven long open reading frames (ORFs) which corresponded to seven known virion proteins were revealed. The complete nucleotide sequence of the Marburg virus genome will be published in a separate article.

The nucleotide sequence of the GP-gene (plus-strand

Correspondence address: A.A. Bukreyev, NPO 'VECTOR', Institute of Molecular Biology, Koltsovo, 633159, Novosibirsk Region, Russian Federation. Fax: (7) (383) 232 8831.

The presented sequence of the Marburg genome fragment has been published in the EMBL Data Library (X68493, 1992).

Abbreviations ARV, avian reticuloendotheliosis virus; ASV, avian sarcoma virus UR2 type; BAEV, baboon endogenous virus; FeLV, feline leukemia virus; HTLV-I, human T-cell leukemia virus type I; M-MuLV, Moloney murine leukemia virus; RSV, Rous sarcoma virus; aa, amino acid.

184

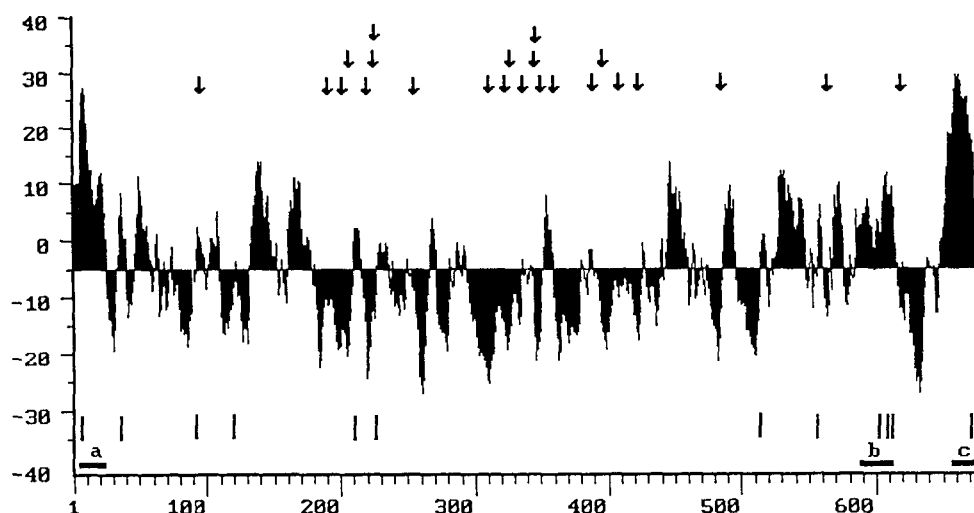


Fig. 2. Schematic representation of the GP-protein based on the predicted amino acid sequence. Hydropathic plot computed by the method of Kyte and Doolittle [12] using an interval of 9 aa. The plotted line at the -5 value represents the midpoint line. The hydrophobic domains are above the line and the hydrophilic domains are below it. The positions of potential glycosylation sites (↓) and the cysteine residues (|) are shown. The predicted signal peptide, 'immunosuppressive domain' and membrane-traversing region are indicated as a, b and c, respectively.

of cDNA), which corresponds to ORF4 [6], and deduced amino acid sequence are presented in Fig. 1. Computer analysis of the full-length RNA sequence revealed the canonic sequences corresponding to those shown in Fig. 1 before and after each ORF. One of these sequences constitutes part of the transcriptional start signal 3'-NNCUNCNUNUAAUU-5' (negative-strand RNA), described in [6] for Marburg Musoke strain and shown to be mRNA extremities. The other sequence corresponds to a transcriptional stop signal and mRNA extremity sequence 3'-UAAUUCUUUUU-5' [6]. The initiation codon of the ORF corresponds to the Kozak rule [11]. The deduced polypeptide is 681 aa long. It has 23 potential asparagine-linked sites of glycosylation (Figs. 1 and 2). Most of these sites are located in two clusters: aa 171–255 and 310–422.

A hydropathic plot [12] (Fig. 2) shows that the amino acid sequence contains two extensive regions of hydrophobic and mostly uncharged amino acids. The first of them is situated at the N-terminal of the protein (aa 6–22). It probably corresponds to the hydrophobic fragment of the signal peptide. The second and most striking of the hydrophobic regions (aa 651–673) is located at the C-terminal of the polypeptide. This fragment is predicted to be a membrane-traversing region of the GP-protein. The last 8 aa at the C-terminal from aa 674, are considered to lie on the inner side of the membrane.

There are two basic residues immediately near the hydrophobic domain: arginine (position 674), and lysine (678). It is possible that any interactions between the protein and the virion core particles occur through these residues. Most of the GP-protein located between two hydrophobic regions has a long hydrophilic region which includes two clusters of potential glycosylation sites and is predicted to lie outside the lipid envelope.

Recently it was reported that the C-terminal part of the GP-protein of Ebola virus has significant homology with the 'immunosuppressive domain' of the PI5E envelope proteins of various oncogenic retroviruses [13]. We have compared the deduced amino acid sequence of GP-protein of Marburg virus with that of the C-terminal of the GP-protein of Ebola virus and with the 'immunosuppressive domain' of RSV, ASV, M-MuLV, FeLV, HTLV-I, ARV, and BAEV [14–20]. Alignment of the C-terminus of Marburg virus GP-protein (from aa 518) with the appropriate region of Ebola virus GP-protein is presented in Fig. 3. Similarity in this region is 49%. This fact confirms the evolutionary relationship between these two members of the Filoviridae family. The comparison of Marburg virus GP-protein (aa 586–611) with that of Ebola virus and with the 'immunosuppressive domain' of oncogenic retroviruses is presented in Fig. 4. The homology between the Marburg and Ebola viruses in this fragment reaches 81%. The align-

←

Fig. 1. The nucleotide sequence of the Marburg virus GP-gene (cDNA plus-strand). The deduced amino acid sequence is given below the nucleotide sequence. The canonic sequences corresponding to transcriptional start and stop signals are overlined. The possible asparagine-linked glycosylation sites are boxed. Two hydrophobic regions located at the N- and C-terminal parts of the putative polypeptide are underlined. The 'immunosuppressive domain' is shaded.

MARBURG	IWSVOEDDLAAGLSWIPFFGPGIEGLYTAGLIKNQNNLVCRRLRLANQTAKSLELL
EBOLA	V TT DEGA I. A V AA. I. IE MH. DG I G Q. E TGA Q F
MARBURG	LRVTTEERTFSLINRHAIDFLLRWGGTCKVLGPDCIGIEDLSRNISEQIDQIKK
EBOLA	. A . L IL K Q HI EPH WTK TDK IH
MARBURG	DE-QKEGTGWGLGGKWWTSWDGVLTLNLGILLLSIAVLIALSCICRIFTKVIK
EBOLA	FVD. TLPDQ DNDN. G- RQWIPA. GVTGV. IAV F KFVF----

Fig. 3. Comparison of the Marburg virus GP-protein (C-terminal region from aa 518) with that of Ebola virus [14]. Gaps are indicated by dashes. Amino acids that are the same in Ebola virus and Marburg virus are replaced by dots.

ment revealed the occurrence of a significant degree of homology (42–46%) between Marburg virus and the P15E protein fragments of retroviruses. There is some experimental evidence showing the participation of this retrovirus domain in the immunosuppressive activity of retroviruses. It was demonstrated that lymphocyte blastogenic responses to mitogens and alloantigens are inhibited by FeLV P15E protein; transformation of human lymphocytes by concanavalin A is also blocked by P15E [21,22]; proliferation of murine cytotoxic T-lymphocyte is inhibited by this protein [23]. M-MuLV P15E inhibits macrophage accumulation at inflammatory foci in mice [24]. Moreover, a synthetic peptide (SKS-17), synthesized to correspond to the region of homology between P15E of various retroviruses, inhibited the proliferation of an interleukin-2-dependent murine cytotoxic T-cell line and alloantigen-stimulated proliferation of murine and human lymphocytes [25]. Taking into account these features of the P15E 'immunosuppressive domain' we consider that the similar region in Marburg virus GP-protein could play an important role in the high pathogenicity of Marburg virus.

We revealed another region of similarity with retroviruses in the transmembrane domain of the GP-protein. It is the sequence of four leucines at position 658–661. There are identical sequences in the transmembrane domain of P15E-proteins of HTLV-1 and BAEV

[18,20]. The existence of regions of similarity between Marburg virus and different retroviruses may be caused by recombination events between ancestors of these viruses.

REFERENCES

- [1] Martini, G. and Siebert, R. (1971) Marburg Virus Disease, Springer, New York.
- [2] Gear, J.S.S., Cassel, G.A., Gear, A.J., Trappler, B., Clausen, L., Meyers, A.M., Kew, M.C., Bothwell, T.H., Sher, R., Miller, G.B., Schneider, J., Koornhoff, H.J., Gomperts, E.D., Isaacson, M. and Gear, J.H.S. (1975) Br. Med. J. 4, 489–493.
- [3] Morbidity and Mortality Weekly Report. CDC, Atlanta, USA (1980) 29, 145–146.
- [4] Brown, F. (1989) Intervirology 30, 181–186.
- [5] Kiley, M.P., Cox, N.J., Elliott, L.H., Sanchez, A., DeFries, R., Buchmeier, M.J., Richman, D.D. and McCormick, J.B. (1988) J. Gen. Virol. 69, 1957–1967.
- [6] Feldmann, H., Muhlberger, E., Randolph, A., Will, C., Kiley, M.P., Sanchez, A. and Klenk, H.-D. (1992) Virus Res. 24, 1–19.
- [7] Feldmann, H., Will, C., Schikore, M., Slenczka, W. and Klenk, H.-D. (1991) Virology 182, 353–356.
- [8] Bukreyev, A.A., Kolichalov, A.A., Volchkov, V.E., Blinov, V.M., Netesov, S.V. and Sandakhchiev, L.S. (1991) Molekulyarnaya Genetika, Mikrobiologiya i Virusologiya 3, 24–30 (in Russian).
- [9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [10] Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.

Marburg	I	N	R	H	A	I	D	F	L	L	T	-	R	W	G	G	T	C	K	V	L	G	P	D	C	C	I		
Ebola	L				K							Q	-															81	
RSV	Q				A							L	-	A	H			H	G			E	D	V	A	G	M	F	46
ASV	Q				A							L	-	A	H			H	G			E	D	I	A	G	M	F	46
M-MuLV	Q			R	G	L						F	L	K	E					L		A	A		K	E	E	F	42
FeLV	Q			R	G	L						F	L	Q	E					L		A	A		K	E	E	F	38
HTLV-I	Q			R	G	L						F	W	E	Q					L			A		Q	E	Q	F	46
ARV	Q			R	G	L						A	E	Q						I		L	A		Q	E	K	F	46
BAEV	Q			R	G	L						A	E	Q						I		L	A		Q	E	K	F	46

Fig. 4. Comparison of Marburg virus GP-protein (aa 586–611) with that of Ebola virus [14] and with the regions of RSV, ASV, M-MuLV, FeLV, HTLV-I, ARV and BAEV P15E protein [15–21]. Dots indicate residues identical with the Marburg virus polypeptide sequences; dashes indicate gaps. Residues identical or belonging to the same group of chemical similarity to Marburg virus are boxed. The amino acids are partitioned into families as follows: (P,G,S,T,A); (F,Y,W,I,L,M,V); (D,E,N,Q); (K,R,H); (C). Percentages of amino acids identical to that of the Marburg virus protein sequence are indicated to the right.

- [11] Kozak, M. (1986) *Cell* 44, 283–292.
- [12] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.
- [13] Volchkov, V.E., Blinov, V.M. and Netesov, S.V. (1992) *FEBS Lett.* 305, 181–184.
- [14] Schwartz, D., Tizard, R. and Gilbert, W. (1983) *Cell* 32, 853–869.
- [15] Neckameyer, W.S. and Wang, L.-H. (1985) *J. Virol.* 53, 879–884.
- [16] Shinnick, T.M., Lerner, R.A. and Sutchffe, J.G. (1981) *Nature* 293, 543–548.
- [17] Elder, J.H. and Mullins, J.I. (1983) *J. Virol.* 46, 871–880.
- [18] Seiki, M., Hattori, S., Hirayama, Y. and Yoshida, M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3618–3622.
- [19] Wilhelmsen, K.C., Eggleton, K. and Temin, H.M. (1984) *J. Virol.* 52, 172–182.
- [20] Kato, S., Matsuo, K., Nishimura, N., Takahashi, N. and Takano, T. (1987) *Jpn. J. Genet.* 62, 127–137.
- [21] Mathes, L.E., Olsen, R.G., Hebebrand, L.C., Hoover, E.A., Schaller, J.P., Adams, P.W. and Nickols, W.S. (1979) *Cancer Res.* 39, 950–955.
- [22] Copelan, E.A., Rinehart, J.J., Lewis, M., Mathes, L., Olsen, R. and Sagone, A. (1983) *J. Immunol.* 131, 2017–2020.
- [23] Orosz, C.G., Zinn, N.E., Olsen, R.G. and Mathes, L.E. (1985) *J. Immunol.* 134, 3396–3403.
- [24] Cianciolo, G.J., Matthews, T.J., Bolognesi, D.P. and Snyderman, R. (1980) *J. Immunol.* 124, 2900–2905.
- [25] Cianciolo, G.J., Copeland, T.D., Oroszlan, S. and Snyderman, R. (1985) *Science* 230, 453–455.